The English translation of Chinese Patent Aplication 02158110.X

A HUMAN HEPATOCELLULAR CARCINOMA-RELATED GENE, ITS ENCODED PRODUCTS AND APPLICATIONS

FIELD OF THE INVENTION

The present invention relates to a gene and its encoded products, as well as their applications in genetic engineering, specially relates to a human hepatocellular carcinoma-related gene and its encoded products, as well as their applications.

BACKGROUND OF THE INVENTION

Hepatocellular carcinoma (HCC) is the major health problem threatening human lives. It is reported that the new cases of primary hepatocellular carcinoma exceeds over one million worldwide each year. 70% of the new cases occur in Asia, and about 40-45% occur in China. The total number of new hepatocellular carcinoma cases every year in China is about 450,000, and the number is increasing. Not only the high incidence, but also difficulty in early diagnosis, fast growing rate, high reoccurrence, and the high mortality rate make HCC a most malignant cancer. Most HCC patients have already progressed to the intermediate stage or late stage when diagnosed, and they can only survive for 3-6 months if without a proper treatment.

To elucidate the mechanism underlying cancerogenesis would help for HCC prevention, diagnosis and treatment. Early diagnosis is crucial for raising the curative rate and reducing the mortality. Currently used HCC-diagnostic marker, the serum AFP, has 30% of negative results in HCC patients, while some benign liver diseases can cause a significant increase of AFP level in serum, creating some difficulty in differential diagnosis.

Cancer is essentially a cellular hereditary disease. Although a great number of HCC-related genes have been discovered, the mechanisms of the hepatocarcinogenesis and its development remain to be elucidated. The known oncogenes can be divided into five categories according to the cellular localization and function of their encoded proteins: I. genes that encode growth factors, including *sis, int-2, hst, fgf-5*; II. genes that encode growth factor receptors, including *erbB, erbB-2, fms, met, ros,* and others; III. genes that encode signal transduction molecules in cytoplasm, including *abl, src, ras, raf, yes, fgr, fes, lck, mos,* and others; IV. genes that encode regulatory molecules for cell proliferation and apoptosis, including *bcl-1, bcl-2* and others; and V. genes that encode the nuclear DNA-binding proteins (transcription factors) , such as *myc, myb, fos, jun, B-lym, ski, ets, rel* and others. It has been demonstrated that *ras, src, myc, met* and *p53* etc. are the genes closely associated with HCC.

SUMMARY OF THE INVENTION

The present invention discloses a novel human HCC-related gene and its encoded products.

This novel human HCC-related gene disclosed by this invention is designated as *LAPTM4B*. It comprises one of the following nucleotide sequences:

1). SEQ ID No: 1 in the sequence listings;

2). Polynucleotides that encode the protein sequences of SEQ ID No: 4, or SEQ ID No: 5, in the sequence listings;

3). DNA sequences having more than 90% homology with the DNA sequences defined by SEQ ID No: 2, or SEQ ID No: 3 in the sequence listings. These DNA sequences encode proteins having the same functions.

SEQ ID No: 1 in the sequence listings contains 954 bases. It is an intact open reading frame. SEQ ID No: 1 has two translational starting sites, the first one is the codon (triplet) at nt1-3 at 5' terminal, and the second one is the codon (triplet) at nt274-276 at 5' terminal. Two full length cDNAs related to SEQ ID No: 1 have two alternative tailing signals. When 5' terminal in SEQ ID No: 1 is extended outward by 85 bases, and 3' terminal is extended outward by 401 bases, SEQ ID No: 2 in the sequence listings is obtained. This expressive sequence of the gene contains 1440 bases. When 5' terminal in SEQ ID No: 1 is extended outward by 85 bases, and 3' terminal is extended outward 1130 bases, SEQ ID No: 3 in the sequence listings is obtained. This expressive sequence of the gene contains 2169 bases. The *LAPTM4B* gene is mapped to chromosome 8q22.1.

The human HCC-related LAPTM4B proteins possesses the amino acid sequence of SEQ ID No:4, or SEQ ID No:5 in the sequence listings. Or it consists of the sequence 4 or the sequence 5, after one or several amino acid residues are replaced, deleted, or added. However, the above altered sequence 4, and/or the sequence 5, still have the same activity to the unchanged sequence 4, and/or the sequence 5.

SEQ ID No: 4 in the sequence listings consists of 317 amino acid residues encoded by the whole sequence of SEQ ID No: 1. Its apparent molecular mass is 35kDa and the putative isoelectric point is 9.05. SEQ ID No: 5 in the sequence listings contains 226 amino acid residues encoded by the segment of bases from nt274th to nt954th in the SEQ ID No: 1. Its apparent molecular mass is 24kDa, and the putative isoelectric point is 4.65.

LAPTM4B proteins in SEQ ID No: 4 and SEQ ID No: 5 have four fragments of membrane-spanning sequences, one potential N-glycosylation site, and a typical lysosome targeting signals. They both belong to the protein superfamily of the tetratransmembrane proteins. However, they have various number of phosphorylation sites. Experiments in the present invention show that LAPTM4B-35 in SEQ ID No: 4 can form a complex in plasma membrane with epidermal growth factor recceptor (EGFR) and integrin α6β1 (the specific receptor of laminin in the extracellular matrix), suggesting the possible function of it on coupling these two proliferation signals from growth factor and extracellular matrix (ECM). These studies further elucidate the molecular mechanism of anchorage-dependent cell growth (i.e. both signals from growth factor and ECM are needed for triggering proliferation) of normal

eukaryotic cells and make a breakthrough progress on interpreting the regulating mechanism of cell proliferation. Experiments in the present invention also demonstrate that the tyrosine residue (Tyr_{285}) in the C terminal of LAPTM4B protein can be phosphorylated. The phosphorylated Tyr_{285} forms a binding site for the SH2 domain of intracellular signal molecules. In the meantime, the N terminal sequences of LAPTM4B contain Pro-rich domains and the typical binding sites for SH3 domain of intracellular signal molecules. The studies mentioned above indicate that SEQ ID No: 4 LAPTM4B protein may be an important docking protein for signal transduction. It can recruit related signal molecules to complete the signal transduction for cell proliferation. Experimental results in the present invention show that the transfection of mouse NIH3T3 cell line by cDNA which encodes the amino acid sequence in SEQ ID No: 4 produces stable transfected and LAPTM4B-35 over expressed NIH3T3-AE cell lines. The proliferation of the transfectants are accelerated and less dependent on serum, also the *LAPTM4B* over-expressed NIH 3T3 cells can form a moderate malignant fibrosarcoma. On the contrary, the transfectant from HLE hepatocellular carcinoma cell line by cDNA, which encodes the amino acid sequence in SEQ ID No: 5 (An encoding sequence where 91 amino acids in the N terminal of LAPTM4B-35 is truncated), cannot form colony; whereas the transfectant from the same cell line by the cDNA, which encodes the amino acid sequence in SEQ ID No:4, can form a larger number of bigger colony. The result shows that the LAPTM4B-35 protein in SEQ ID No: 4 and LAPTM4B-24 protein in SEQ ID No: 5 have reciprocally antagonistical functions in cell proliferation and survival.

LAPTM4B gene is expressed at low level in normal liver tissues. However, its expression in 96.88% (31/32) human hepatocellular carcinoma tissues is significantly up-regulated, and slightly up-regulated in some of paired non-cancerous liver tissues. In vast majority of cell lines from hepatoma tissues tested are highly expressed. In addition, *LAPTM4B* is expressed to various extent in some other cancer cell lines and some human normal tissues, such as testes, heart and skeleton muscle. The expression vectors containing sequences described in SEQ ID No: 1, 2, 3, and the transfected cell lines containing SEQ ID No: 1, 2, 3 sequences are all encompassed by the present invention.

The LAPTM4B proteins disclosed by the present invention can be used as novel early diagnostic markers for HCC, The LAPTM4B-specific ELISA and the related kit can improve the early diagnostic rate and the accurate of diagnosis for HCC. *LAPTM4B* gene and its encoded proteins can also be the cancer target to develop novel anti-cancer strategies, for example using RNAi (or siRNA) technique to silence the expression of *LAPTM4B*; and LAPTM4B protein may be used as targets to exploit novel anti-cancer drugs. Thus, LAPTM4B gene and its encoded proteins as targets would create novel anti-HCC approaches and would make significant impacts on human society.

Further illustrations are made in the following description and figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows profiles from Northern Blot analysis

Figure 2 shows the immunocytochemical diagram.

Figure 3 shows the profiles from Western Blot analysis.

Figure 4 also shows a profile from Western Blot analysis.

Figure 5 shows cell growth curves.

Figure 6 also shows cell growth curves.

Figure 7 shows the oncogenic effect on mouse of cDNA-transfected cells generated in the present invention.

Figure 8 is a histogram showing the level of the LAPTM4B antigen in the serum of patients with hepatocellular carcinoma.

DETAILED DESCRIPTION OF THE INVENTION:

Example 1: Northern Blot analysis of *LAPTM4B* expressions in four types of liver tissues at various proliferation and differentiation status.

Four types of liver tissues at various proliferation and differentiation status were chosen. They were from normal adult livers (NL, with little proliferation and high differentiation), fetal livers (FL, at active proliferation and low differentiation), hepatocellular carcinoma (HCC, uncontrolled proliferation and abnormal differentiation), and paired non-cancerous livers (PNL, generally is of precancerous stage in an active proliferation status). The Northern Blot analysis was used to detect the transcription of gene in these tissues. RNA samples were extracted from 5 normal adult liver tissues, 32 HCC tissues, and 32 paired non-cancerous liver tissues freshly obtained from surgical excision, and 5 fetal liver tissues from abortive fetus, respectively. After electrophoretic separation, they were transferred to a nylon film and hybridized by Dig-labeled LAPTM4B probe. The film was washed at 68°C and the hybridization signals were developed according to the manual. The results are shown in Figure 1. Lane 1 represents the expression of fetus livers. Lane 2 represents the expression of normal adult liver. Lanes 3, 5, 7, and 9 represents the expression of HCC tissues. Lanes 4, 6, 8, 10 represents the expression of PNL tissues. The results show that the expression of *LAPTM4B* in various liver tissues has the following order: HCC tissue > PNL tissue and fetal liver tissue > normal adult liver tissue.

Example 2: Clonings of *LAPTM4B*

By using fluorescent differential-display PCR technique, a cDNA segment (LC27) of an unknown gene was screened out from the differential display spectra of four types of human liver tissues in different proliferation and differentiation status, such as normal adult livers (NL), fetus livers (FL), cancerous livers (HCC), and paired non-cancerous liver (PNL). The LC27 segment (426bp) was elongated towards the 5' direction first by splicing homogenous sequences according to the EST, and then by RACE (rapid amplification of cDNA ends) with high temperature RT-PCR techniques. Two full-length cDNA sequences, i.e., SEQ ID No. 2

and 3 in the sequence listing, were obtained after sequencing.

Example 3: LAPTM4B protein localization analyzed by immunocytochemical analysis

The HCC cells from BEL-7402 cell line was fixed by 4% polyformaldehyde on to a slide and stayed at RT for 30 min. After washing by PBS, the slide was put into 3% H_2O_2 in methanol solution for 15min to block endogenous peroxidase activities and then washed with 0.05% Tween 20/PBS. After incubating in 10%BSA at 37°C for 30min, specific antibody (anti-LAPTM4B-EC2-pAb, 1:100) was added onto the slide and incubated at 4°C overnight. After washing with PBS, the slide was further incubated with the second anti-rabbet antibody from goat at 37°C for 30min. After washing with PBS, the sample was stained with DAB, the nuclear was then stained by Meyer's Hematoxylin, The result (Fig.2) shows that LAPTM4B localizes at membranous structures, including endomembranes and likely plasma membrane.

Example 4: Western Blot analysis of LAPTM4B protein in subcellular fractions

The HCC, PNL and NL tissues were homogenized in 150 mM NaCl, 10mM Tris buffer (pH7.4), 0.5% Triton X-100, 5mM EDTA, 1mM PMSF. Then centrifuged by differential centrifugation, the nuclei, mitochondria and membrane fractions were separated. Then SDS-PAGE of these 3 fractions was performed. The immunobloting was performed with polyclonal anti-LAPTM4B antibody. The results (Fig. 3) shows that LAPTM4B protein existed at membrane fractions, but not nuclear and mitochondria. N: nuclear, M: membranes, Mit: mitochondria; Lanes 1,2,3 from HCC; Lane 4,5,6 from PNL; Lanes 7,8,9 from NL.

Example 5: Western Blot analysis of partitional products of liver tissues by Triton X114

The HCC, PNL, and NL tissues from surgical resection were fully homogenized in a lysis buffer (150mM NaCl, 10mM Tris buffer pH7.4, 1%Triton X-114, 5mM EDTA, 1mM PMSF) and centrifuged to remove the debris. 300μl of TBS containing 6% sucrose, 0.06%triton X-114 was added to 200μl of the cleared supernatant and incubated at 37°C for 5min. After centrifugation the upper layer (aqueous phase) and lower layer (detergent phase) was separated. 1%Triton-X114 was added to the aqueous phase, and put onto ice for 10min, then incubated at 37°C for 5min. After centrifugation the upper layer (aqueous phase) was collected, and combined with the former aqueous phase. This step was repeated once. The aqueous phase and the detergent phase was washed with TBS containing 2% Triton X-114 twice, respectively. Then twice volume cold ethanol was added to precipitate proteins. After centrifugation the protein precipitant was collected and dissolved into SDS-PAGE loading buffer. After electrophoresis the proteins were blotted with anti-LAPTM4B antibody. The result (Fig.4) shows that LAPTM4B proteins all exists in membranous fraction (the detergent phase),

but not in aqueous phase. Lane A: homogenate before partition; Lane B: the aqueous phase after partition; Lane C: the detergent phase after partition.

Example 6: Regulatory effect of the gene disclosed in the present invention on cell proliferation as demonstrated by a full-length cDNA transfection.

The *pcDNA 3.0* vector was used to construct the expressive plasmids containing the full length *LAPTM4B* cDNA sequence (2.16kb). mouse NIH3T3 cells were transfected by the plasmids and the stably high expressed transfectants were screened and cloned. The cell growth curves were measured by detecting the viable cells number determined by acidic phosphatase activity; the cell cycle was analyzed by the flow cytometry; the expression level of cyclin E was detected by Western Blot analysis. The results (Fig.5) show that the proliferation of NIH3T3 cells was accelerated by transfection of full length *LAPTM4B* cDNA, and the expression of cyclin E in the *LAPTM4B*-transfectants is remarkably enhanced; In the meantime, the proliferation of *LAPTM4B*-transfectants is less dependent on the growth factors in serum (Fig.6). All the results indicate that this gene may participate in the regulation of cell proliferation and its over expression (activation) may relate to the dysregulation of cell proliferation.

Example 7: Tumorigenic effect of *LAPTM4B cDNA*-transfected cells on mouse.

Six-weeks-old male mice were randomly divided into three groups: In the first control group, the mice were injected with physiological saline as a control. In the second control group, the mice were inoculated with the *pcDNA3* MOCK (no-load plasmid) transfected cells. In the test group, all the mice were inoculated with the *LAPTM4B* cDNA-transfected cells. Each mouse was subcutaneously inoculated with $2x10^6$ cells. There were four to six mice in each group. The mice were sacrificed after 21 days inoculation and dissected. The results are shown in Fig.7. Two mice (half of inoculated mice) in the test group developed a clearly moderate malignant fibrosarcoma (A, B); the other two mice were identified as liquefied lymphoid tissue at the inoculated sites (C, D). In contrast, twelve mice in the two control groups showed no sign of tumor formation untill being inoculated for 86 days.

The data in Examples 6 and 7, as well as the expressive spectra indicate that *LAPTM4B* may be a novel proto-oncogene.

Example 8: Primary analysis of LAPTM4B antigen in the serum of patients with

hepatocellular carcinoma by the ELISA method

96 wells culture plates were coated with sera in various dilutions from HCC patients or normal individuals at 4°C overnight. Each well was washed with 0.5% Tween-20 in PBS washing solution, and then 2% BSA was added for blocking at room temperature for 1 hour. Then LAPTM4B-10 peptides-pAb antibody in various dilutions was added and incubated at room temperature for 2 hours. After washing with PBS the goat anti-rabbit antibody labeled by horseradish peroxidase (1:1000 times dilution) was added. After standing at room temperature for 2 hours and PBS washing, 1 μ g/mL o-phenyldiamine solution was added and incubated for 10-15 minutes to develop color and then H_2SO_4 was used to stop the reaction. The microtiter was used to measure the OD at 490 nm and the antigen level was estimated. The result (Fig.8) shows that the sera of patients with hepatocellular carcinoma contains higher level of LAPTM4B antigen than that from normal individuals, indicating that LAPTM4B has a potential to become a new marker for hepatocellular carcinoma diagnosis.